Staff Papers Series

P 69-1 January 1969

TESTING HYPOTHESIS ON THE RELATIVE SIZE OF THE COEFFICIENTS IN REGRESSION MODELS

J. S. Mann

Department of Agricultural Economics

University of Minnesota
Institute of Agriculture
St. Paul, Minnesota 55101

P 69-1 January 1969

TESTING HYPOTHESIS ON THE RELATIVE SIZE OF THE COEFFICIENTS IN REGRESSION MODELS

J. S. Mann

Testing Hypothesis on the Relative Size of the Coefficients in Regression Models

J. S. Mann

In the traditional multiple regression analysis a hypothesis about a set of regression coefficients is tested. This hypothesis is generally stated in terms of one or more coefficients being equal to zero or some other given value. (See, for example, Johnston [3, Section 4.3], Goldberger [1, Section 7, Chapter 4] and Mood and Graybill [4, Theorem 13.6]). However, often the researcher is faced with the problem of testing a hypothesis about the relative size of various coefficients; for example, two coefficients are equal. This can be handled with the help of a matrix of restrictions (r in the following discussion). Johnston [3, pp. 131-133] has developed a test for a simple case where r has rank 1. In this paper a readily usable technique for the general case is presented. First, the traditional regression model is set up. Second, the general method for testing hypothesis on the relative size of the coefficients is developed. Finally, the method is applied to a few leading cases.

The Regression Model

Let y be a linear function of k independent variables x_1, x_2 --- x_k and a disturbance term u. If a sample of T observations on y and x's is taken, the relationship can be written as

$$y_t = \beta_o + \beta_1 x_{1t} + \beta_2 x_{2t} + --- + \beta_k x_{kt} + u_t$$

for t = 1, 2, --, T

These T equations can be set out in matrix notation as

$$y = X \beta + u$$

where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_t \end{bmatrix},$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & --- & x_{k1} \\ 1 & x_{12} & x_{22} & --- & x_{k2} \\ . \\ . \\ . \\ 1 & x_{1T} & x_{2T} & --- & x_{kT} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_o \\ \beta_1 \\ . \\ . \\ . \\ \beta_k \end{bmatrix}$$

and

$$u = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_T \end{bmatrix}$$

If we make the following assumptions:

(1) u is normally distributed with Eu = 0 and $Euu' = \sigma^2 I$,

(2) X has rank k+1

then, the maximum likelihood estimator of β is [2, p.111]

$$\hat{\beta} = (X' X)^{-1} X' y$$

General Hypothesis

The general hypothesis can be stated as

$$H_o : r \beta = R$$

where r is a G x k + 1 matrix of G independent restrictions on β (that is, r has rank G). For example, consider a regression problem with three independent variables

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t$$

We want to test $\beta_2 = \beta_3$. The null hypothesis can be stated as

$$H_o : \beta_2 = \beta_3$$

or

$$\beta_2 - \beta_3 = 0$$

or

$$(0 \; 0 \; 1 \; -1) \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = 0$$

Here r = (0 0 1 -1) and R = 0. (In this case G = 1.)

As another example, in a three independent variables case, let the null hypothesis be $\beta_0 = 0$ <u>and</u> $\beta_2 = \beta_3$. In other words,

$$H_o : \beta_0 = 0 \text{ and}$$

$$\beta_2 = \beta_3$$

or

$$\beta_0 = 0 \text{ and}$$

$$\beta_2 - \beta_3 = 0$$

or

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Here

$$r = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix},$$

$$R = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and G = 2.

The general hypothesis $r\beta = R$ can also be written as

$$\delta = (r\beta - R) = 0 \qquad (1)$$

If the estimated value of β, that is $\hat{\beta}$, is substituted, we obtain

$$\hat{\delta} = (r\hat{\beta} - R) \qquad (2)$$

Subtracting (1) from (2)

$$\hat{\delta} - \delta = r(\hat{\beta} - \beta)$$

If the null hypothesis is true, that is $\delta = 0$,

$$\hat{\delta} = r(\hat{\beta} - \beta).$$

$\hat{\delta}$ is an indication of how much the estimated value $r\hat{\beta}$ deviates from the true value $r\beta$.

Now we proceed to derive a test statistic for $\hat{\delta}$. It is known that $\hat{\beta} - \beta$ has normal distribution with mean 0 and variance $\sigma^2 (X' X)^{-1}$ [4, Theorem 13.3]. Since a linear function of a normally distributed random variable has a normal distribution, $\hat{\delta} = r(\hat{\beta} - \beta)$ is normally distributed with

$$E\hat{\delta} = rE(\hat{\beta} - \beta) = 0, \text{ and}$$

$$E\hat{\delta}\hat{\delta}' = rE(\hat{\beta} - \beta)(\hat{\beta} - \beta)'r'$$

$$= \sigma^2 r(X' X)^{-1} r'.$$

Using the fact that if a_i be independent normally distributed variables (i = 1, 2, --, k) with mean = 0 and variance $= \sigma^2$, then

$$z = \frac{\Sigma a_i^2}{\sigma^2}$$

has a χ^2 distribution with k degrees of freedom, in the present case

$$\frac{\hat{\delta}' [r(X'X)^{-1} r']^{-1} \hat{\delta}}{\sigma^2}$$

has χ^2 distribution with G degrees of freedom.

Further, $\frac{\hat{u}' \hat{u}}{\sigma^2}$ is also distributed as χ^2 with T-k -1 degrees of freedom and $\hat{u}' \hat{u}$ is independent of $\hat{\delta}$ [4, Theorem 13.3]. The ratio of two independent Chi-square variables each divided by its degrees of freedom has the Snedecor F distribution. Therefore, the ratio

$$\frac{\frac{1}{\sigma^2} \frac{1}{G} \hat{\delta}' [r(X' X)^{-1} r']^{-1} \hat{\delta}}{\frac{1}{\sigma^2} \frac{1}{T-k-1} \hat{u}' \hat{u}}$$

$$= \frac{(T-k-1)\,[r\,(\hat{\beta}-\beta)]'\,[r\,(X'X)^{-1}\,r]^{-1}\,[r\,(\hat{\beta}-\beta)]}{G\,\hat{u}'\,\hat{u}}$$

has $F_{G,\,T-k-1}$ distribution.

Applications

The discussion in this section is limited to a case of three independent variables:

$$\hat{y}_t = \hat{\beta}_o + \hat{\beta}_1\, x_{1t} + \hat{\beta}_2\, x_{2t} + \hat{\beta}_3\, x_{3t},$$

where $\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = (X'X)^{-1}\, X'y.$

Let the i jth element of $(X'X)^{-1}$ be designated as c_{ij}

Case 1. $H_o:\ \beta_2 = \beta_3$

An economic example would be the coefficients to two related goods in the demand function for a commodity. It may be conjectured that those two goods have identical influence on the demand for the given commodity and may be combined into one good. In this case the hypothesis of equality of two coefficients is tested.

The null hypothesis can also be stated as

$$H_o:\ \beta_2 - \beta_3 = 0,$$

or

$$(0\ 0\ 1\ -1)\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{bmatrix} = 0$$

Thus, $r = (0\ 0\ 1\ -1)$ and $R = 0$. Hence $\hat{\delta} = (\hat{\beta}_2 - \hat{\beta}_3)$.

Then

$$r\ (X'X)^{-1}\ r' = (0\ 0\ 1\ -1)\begin{bmatrix}c_{00} & c_{01} & c_{02} & c_{03}\\ c_{10} & c_{11} & c_{12} & c_{13}\\ c_{20} & c_{21} & c_{22} & c_{23}\\ c_{30} & c_{31} & c_{32} & c_{33}\end{bmatrix}\begin{bmatrix}0\\ 0\\ 1\\ -1\end{bmatrix}$$

$$= (c_{20}-c_{30}\quad c_{21}-c_{31}\quad c_{22}-c_{32}\quad c_{23}-c_{33})\begin{bmatrix}0\\ 0\\ 1\\ -1\end{bmatrix}$$

$$= c_{22} - c_{32} - c_{23} + c_{33}$$

The test statistic is

$$F_{1,T-4} = \frac{(T-4)\ (\hat{\beta}_2 - \hat{\beta}_3)\ [c_{22} - c_{32} - c_{23} + c_{33}]^{-1}(\hat{\beta}_2 - \hat{\beta}_3)}{\hat{u}'\ \hat{u}}$$

Case 2: We may want to test the hypothesis that two coefficients are in a given ratio. For example,

$$H_o:\ \beta_2 = k\ \beta_3$$

$$\text{or}\quad \beta_2 - k\ \beta_3 = 0$$

$$\text{or}\ (0\ 0\ 1\ -k)\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{bmatrix} = 0$$

Here, $r = (0\ 0\ 1\ -k)$ and $R = 0$.

Then

$$r\,(X'\,X)^{-1}\,r' = (0\;0\;1\;-k)\begin{bmatrix} c_{00} & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \\ c_{30} & c_{31} & c_{32} & c_{33} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \\ -k \end{bmatrix}$$

$$= (c_{20} - kc_{30} \quad c_{21} - kc_{31} \quad c_{22} - kc_{32} \quad c_{23} - kc_{33})\begin{bmatrix} 0 \\ 0 \\ 1 \\ -k \end{bmatrix}$$

$$= (c_{22} - kc_{32} - kc_{23} + k^2 c_{33})$$

and

$$F_{1,T-4} = \frac{(T-4)\;(\hat{\beta}_2 - k\hat{\beta}_3)\;[c_{22}-kc_{32}-kc_{23}+k^2c_{33}]^{-1}(\hat{\beta}_2-k\hat{\beta}_3)}{\hat{u}'\;\hat{u}}$$

is the test statistic.

Case 3: Consider a Cobb - Douglas production function

$$y_t = \alpha\, x_{1t}^{\beta_1}\, x_{2t}^{\beta_2}\, x_{3t}^{\beta_3}\, u_t\,.$$

Taking log of both sides

$$\log y_t = \beta_0 + \beta_1 \log x_{1t} + \beta_2 \log x_{2t} + \beta_3 \log x_{3t} + \log u_t$$

where $\beta_0 = \log\alpha$

We want to test the hypothesis of constant returns to scale, that is

$$H_o:\; \beta_1 + \beta_2 + \beta_3 = 1$$

$$\text{or } (0\;1\;1\;1)\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = 1$$

Here, r = (0 1 1 1) and R = 1

Then

$$r\,(X'X)^{-1}r' \;=\; (0\;1\;1\;1)\begin{bmatrix} c_{00} & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \\ c_{30} & c_{31} & c_{32} & c_{33} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= (c_{10} + c_{20} + c_{30} \quad c_{11} + c_{21}\; c_{31} \quad c_{12} + c_{22} + c_{32} \quad c_{13} + c_{23} + c_{33})\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= (c_{11} + c_{21} + c_{31} + c_{12} + c_{22} + c_{32} + c_{13} + c_{23} + c_{33})$$

(Here X'X is the matrix of sums of squares and cross products of the logs of x variables).

The test statistic is

$$F_{1,\,T-4} = \frac{(T-4)\;(\hat\beta_1+\hat\beta_2+\hat\beta_3-1)\,[\,c_{11}+c_{21}+c_{31}+c_{12}+c_{22}+c_{32}+c_{13}+c_{23}+c_{33}\,]^{-1}(\hat\beta_1+\hat\beta_2+\hat\beta_3-1)}{\hat u'\,\hat u}$$

<u>Case 4:</u> In order to demonstrate the method for more than one restriction, consider a demand curve

$$y_t = \alpha\; x_{1t}^{\beta 1}\;\; x_{2t}^{\beta 2}\;\; x_{3t}^{\beta}\;\; u_t$$

where,

y_t = per capita quantity demanded,

x_{1t} = price of the commodity,

x_{2t} = price of the competing good, and

x_{3t} = per capita income.

The function can be fitted by taking log of both sides and following the usual procedure of solving the normal equations, X'X being in terms of logs of the x variables.

We may want to test that the commodity has unit price elasticity of demand <u>and</u> that the demand function satisfies the homogeneity condition, that is the sum of own price elasticity, cross elasticity and income elasticity is zero.

$$H_o: \beta_1 = -1 \text{ and}$$

$$\beta_1 + \beta_2 + \beta_3 = 0$$

$$\text{or} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

where $\beta_0 = \log \alpha$

Here

$$r = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix},$$

$$R = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

and G = 2.

Now,

$$r\ (X'X)^{-1}\ r' = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_{00} & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \\ c_{30} & c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} c_{10} & c_{11} & c_{12} & c_{13} \\ c_{10}+c_{20}+c_{30} & c_{11}+c_{21}+c_{31} & c_{12}+c_{22}+c_{32} & c_{13}+c_{23}+c_{33} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} c_{11} & c_{11} + c_{12} + c_{13} \\ c_{11}+c_{21}+c_{31} & c+_{11}+c_{21}+c_{31}+c_{12}+c_{22}+c_{32}+c_{13}+c_{23}+c_{33} \end{bmatrix}$$

And the test statistic is

$$F_{2,T-4} = (T-4)\ \frac{[\hat{\beta}_1+1 \quad \hat{\beta}_1+\hat{\beta}_2+\hat{\beta}_3] \begin{bmatrix} c_{11} & c_{11} + c_{12} + c_{13} \\ c_{11}+c_{21}+c_{31} & c_{11}+c_{21}+c_{31}+c_{12}+c_{22}+c_{32}+c_{13}+c_{23}+c_{33} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\beta}_1+1 \\ \hat{\beta}_1+\hat{\beta}_2+\hat{\beta}_3 \end{bmatrix}}{2\,\hat{u}'\ \hat{u}}$$

Finally, it can be shown that the traditional test on a set of coefficients is a special case of the generalized hypothesis. For example,

$$H_o: \quad \beta_2 = 0$$

is equivalent to

$$(0\ 0\ 1\ 0)\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{bmatrix} = 0$$

that is, r = (0 0 1 0).

Now

$$r\ (X'X)^{-1}\ r' = (0\ 0\ 1\ 0)\begin{bmatrix}c_{00} & c_{01} & c_{02} & c_{03}\\ c_{10} & c_{11} & c_{12} & c_{13}\\ c_{20} & c_{21} & c_{22} & c_{23}\\ c_{30} & c_{31} & c_{32} & c_{33}\end{bmatrix}\begin{bmatrix}0\\ 0\\ 1\\ 0\end{bmatrix} = c_{22}$$

And

$$F_{1,T-4} = \frac{(T-4)\ \hat{\beta}_2\ c_{22}^{-1}\ \hat{\beta}_2}{\hat{u}'\ \hat{u}}$$

is the test statistic

This may also be written as

$$F_{1,T-4} = \frac{\hat{\beta}_2^2}{c_{22}\dfrac{\hat{u}'\ \hat{u}}{T-4}} = \left[\frac{\hat{\beta}_2}{\sqrt{c_{22}\hat{u}'\ \hat{u}\ /\ T-4}}\right]^2 = t^2_{T-4}$$

Literature Cited

[1] Goldberger, Authur S. Econometric Theory, New York, John Wiley & Sons, 1964.

[2] Graybill, Franklin A. An Introduction to Linear Statistical Models, Vol I, New York, McGraw-Hill Book Co., 1961.

[3] Johnston, J. Econometric Methods, New York, McGraw-Hill Book Co., 1963.

[4] Mood, Alexander M. and Franklin A. Graybill, Introduction to the Theory of Statistics, New York, McGraw-Hill Book Co., 1963.